# Business data science

Prof. Sergio Venturini

***OBIETTIVO DEL CORSO E RISULTATI DI APPRENDIMENTO ATTESI***

La crescente quantità di dati oggi disponibile rappresenta una delle più grandi rivoluzioni nella recente storia dell’uomo. Una parte essenziale di questa rivoluzione è rappresentata dalla trasformazione di queste enormi moli di dati in preziose intuizioni, che vengono utilizzate dai *decision-maker* a qualsiasi livello per identificare la strategia migliore da intraprendere.

Questo corso si pone essenzialmente due obiettivi: da una parte fornire allo studente gli strumenti analitico-statistici oggi più frequentemente utilizzati nel business sia digitale che tradizionale. Dall’altra parte, il corso mira a favorire nello studente la formazione di una capacità di modellazione della realtà, necessaria per l’analisi quantitativa di fenomeni economici e sociali. La presentazione dei vari argomenti è illustrata attraverso la discussione pratica di casi relativi a diverse applicazioni aziendali come ad esempio la *churn analysis* nell’ambito del customer relationship management, il *credit scoring* nell’ambito del risk management o il *predictive maintenance* nel manufacturing. In parallelo alla presentazione metodologica, il corso si propone anche di mostrare l’impiego di R, uno dei software attualmente più utilizzati per lo sviluppo di applicazioni della *data science* nel business.

***RISULTATI DI APPRENDIMENTO ATTESI***

CONOSCENZA E CAPACITÀ DI COMPRENSIONE

Al termine dell’insegnamento, lo studente sarà in grado di:

* Identificare la metodologia corretta da utilizzare per risolvere il problema oggetto di studio.
* Svolgere le analisi presentate nel corso con il software R.
* Riconoscere e interpretare correttamente i risultati di analisi statistiche applicate a problemi aziendali.

CONOSCENZA E CAPACITÀ DI COMPRENSIONE APPLICATE

Al termine dell’insegnamento, lo studente sarà in grado di:

* Sintetizzare in modo appropriato un insieme di dati mediante grafici e indici.
* Confrontare metodi alternativi di analisi e individuare il metodo migliore per il problema in esame.
* Creare report analitici per diverse applicazioni aziendali, principalmente in ambito marketing.
* Interpretare correttamente i risultati di analisi prodotte autonomamente o da terzi.
* Utilizzare il software R per l’applicazione dei metodi presentati.

AUTONOMIA DI GIUDIZIO

* Apprendimento dei concetti statistici indispensabili per lavorare autonomamente nella ricerca, selezione ed elaborazione dei dati aziendali e nello sviluppo di report analitici.

ABILITÀ COMUNICATIVE

* Apprendimento della terminologia e di metodologie statistiche avanzate indispensabili per implementare e comunicare in modo appropriato i risultati delle analisi condotte in diversi contesti aziendali.

***PROGRAMMA DEL CORSO***

 Il corso si articola nei seguenti argomenti:

1. Review di *descriptive analytics*.
* Sintesi dei dati mediante l’uso di visualizzazioni grafiche.
* Sintesi dei dati mediante l’uso di indici numerici.
1. Review di statistica inferenziale.
* Variabilità campionaria. Il concetto di distribuzione campionaria.
* Stima puntuale e per intervallo per alcuni casi notevoli.
* Introduzione alla verifica di ipotesi e presentazione di alcuni casi notevoli. P-value di un test.
1. Riduzione della complessità di un data set: l’analisi delle componenti principali.
2. Principali metodi di *predictive analytics*.
* Il modello di regressione lineare. Esempi (previsione dei comportamenti di acquisto).
* Il modello di regressione logistica per risposta binaria. Esempi (modelli di *credit scoring* e di *churn*).

***BIBLIOGRAFIA***

 Il materiale corso sarà fornito dal docente attraverso la piattaforma BlackBoard.

Per ulteriori approfondimenti si consiglia di consultare i seguenti testi:

* Chapman, C., McDonnell Feit, E., *R for Marketing Research and Analytics*. Secondaedizione. Springer-Verlag, 2019.
* James, G., Witten, D., Hastie, T., Tibshirani, R., *An Introduction to Statistical Learning*. Seconda edizione. Springer-Verlag, 2021.
* Kuhn, M., Johnson, K. *Applied Predictive Modeling*. Springer-Verlag, 2013.

***DIDATTICA DEL CORSO***

Lezioni frontali

***METODO E CRITERI DI VALUTAZIONE***

L’esame è scritto e consente di ottenere un punteggio massimo di 31/30. Il testo dell’esame potrà contenere sia esercizi pratici sia domande teoriche (definizioni) secondo due possibili modalità: 1) domande a risposta multipla e 2) domande con risposta aperta.

L’esame si riterrà superato se il voto finale sarà maggiore o uguale a 18/30. Un punteggio finale di 31/30 equivale ad un voto di 30 e lode.

L’esame riguarderà tutto il materiale presentato durante il corso.

Gli studenti apprendono il relativo voto tramite i canali istituzionali. Trascorsi 5 giorni dalla comunicazione, il voto dell’esame si intende accettato e sarà registrato d’ufficio. In caso contrario, entro lo stesso termine lo studente dovrà esplicitamente manifestare la propria intenzione di rifiutare il voto attraverso le procedure stabilite dall’Ateneo.

Per ciascuna prova verrà comunicata, insieme al voto conseguito, un’unica data per prendere visione del compito.

Lo studente ha diritto a sostenere l’esame nella modalità sopra descritta **al massimo per tre appelli in un anno accademico**. Per i tentativi successivi al terzo (ovvero dal quarto in poi), l’esame dovrà essere svolto **esclusivamente in forma orale**.

Le modalità d’esame sopra dettagliate mirano a verificare:

* La capacità di identificare la metodologia corretta per risolvere un dato problema.
* La comprensione della logica sottostante una determinata procedura.
* La capacità di calcolare specifici indicatori statistici con il software.
* La capacità di interpretare l’output di analisi ottenute con il software.

***AVVERTENZE E PREREQUISITI***

1. Le regole sopra descritte si applicano indistintamente a tutti gli studenti, compresi quelli in debito d’esame.
2. Il corso presuppone che lo studente abbia dimestichezza con alcuni strumenti matematici di base (potenze, logaritmi, concetto di funzione).
3. Nonostante non sia necessaria, durante gli esami è comunque ammesso l'utilizzo di una calcolatrice (non programmabile).
4. A causa dell’emergenza per COVID-19, le effettive modalità d’esame potrebbero subire variazioni.

***ORARIO E LUOGO DI RICEVIMENTO DEGLI STUDENTI***

Gli orari di ricevimento sono disponibili on line nella pagina personale del docente, consultabile al sito <http://docenti.unicatt.it/>